

METHOD FOR SUPPORTING SDH/SONET APS ON ETHERNET

RELATED APPLICATION

5

This application claims priority to co-pending United States Provisional Application No. 60/317,035, filed September 4, 2001, for all subject matter common to both applications. The disclosure of said provisional application is hereby incorporated by reference in its entirety.

10

FIELD OF THE INVENTION

The present invention relates generally to network switching architecture and more specifically to supporting SDH/SONET Automatic Protection Switching (APS) functionality in an Ethernet network.

15

BACKGROUND OF THE INVENTION

SDH/SONET (Synchronous Digital Hierarchy/Synchronous Optical Network) standards evolved originally for use in a voice network. SDH is a European version of a standard that is substantially the same as the SONET standard developed in North America. SDH/SONET contains connection oriented synchronous TDM circuit switching technology. The SDH/SONET configured network runs at the same clock domain (e.g., every section of the network can be traced to a primary clock reference). The network allocates fixed bandwidth time slots for each circuit. The SDH/SONET architectures are connection based protocols in that there is a physical circuit arrangement between ports in a switch to establish an end to end path. The digital transitions in signals traveling through an SDH/SONET network occur at the same rate, however there may be a phase difference between the transitions of any two signals caused by time delays or jitter in the transmission system.

20

25

30

Ethernet evolved primarily as a data network. In contrast to SDH/SONET, Ethernet is a connectionless asynchronous Carrier Sense, Multiple Access with Collision Detection (CSMA/CD) packet switching technology. The Ethernet architecture does not

rely on a single clock domain like the SDH/SONET architecture. The Ethernet architecture sends a series of packets across the network containing data. Whenever a packet needs to be sent, the transmitter will try to transmit the packet. The Ethernet architecture is also connectionless in that the packets travel from node to node within the
5 network without establishing a logical or physical circuit. The end to end path is discovered through a process called "Bridging". Ethernet is fundamentally a Local Area Networking (LAN) technology.

SDH/SONET networks provide reliable, guaranteed available bandwidth, low
10 jitter connections. These characteristics are required for voice quality networks. SDH/SONET, however, is bandwidth inefficient and has a higher overhead than many other network architectures. Ethernet networks, in contrast, provide lower reliability best effort delivery, and low cost bandwidth connections. These characteristics are suitable for data quality networks. Ethernet has non-guaranteed transmission and low
15 overhead and supports fewer operational functions than SDH/SONET. In SDH/SONET, once the circuit is established, bandwidth is allocated for an application and cannot be used by any other application, even if the original application is not using the bandwidth. In Ethernet, applications only use bandwidth when they need the bandwidth to transmit packets.

20

In SDH/SONET networks, Automatic Protection Switching (APS) functionality is known. SDH/SONET standards define APS controller as the "part of a node that is responsible for generating and terminating information carried in the APS protocol and implementing the APS algorithm." SDH/SONET standards also define APS signaling
25 protocol and APS (K1/K2) bytes. SDH/SONET standards also define various algorithms for linear, ring and mesh protection. SDH/SONET APS functionality can support 50ms switchover, unidirectional and bi-directional switchover, revertive and non-revertive switchover, manual or automatic switchover. SDH/SONET APS functionality can also support linear, ring, and mesh topologies, and Line and Path
30 protections. The APS feature enables the switchover of circuits in case of circuit failure and is often utilized in optical network systems. In general, the APS feature organizes a network into a collection of "working" interfaces and "protect" interfaces. When a working interface fails, a protect interface immediately assumes the working interface

traffic load. In APS there is a working port/link and a protect port/link. Upon initialization and full functioning of a network system, the working port/link is active and the protect port/link maintains a standby mode. If there is an equipment failure during operation, the protect port/link becomes the active port/link, taking over for the failed working port/link, i.e., the protect port/link becomes the new working port/link. Under known APS systems, there can be a minimal traffic disruption during the switchover, on the order of less than 50 ms.

In voice networks, SDH/SONET APS Standard functionality provides the architecture for protection in under 50 ms from equipment failure for ring, linear, or mesh topologies. In order for data networks to be able to support voice traffic, the network must be able to provide the same level of protection both in terms of time to recover and working with different network topologies, i.e., support rings and linear topologies. Ethernet is the most common data network data link layer protocol. There is no Ethernet standard to provide APS functionality.

In Ethernet networks, several standards and proprietary technologies support link failure. Spanning Tree Protocol (STP) IEEE 802.1D standard provides topology changes. STP calculates and maintains the topology by sending and listening to Configuration Messages and several timers. These Configuration Messages are emitted every time a "Hello Timer" times out. Typical this is set to 2 seconds. This means that STP cannot support 50 ms recovery as required for link APS SDH/SONET standard. As the number of nodes grows larger in a STP domain, STP convergence also slows down considerably. It can take minutes to converge. Because of polling, STP also consumes some bandwidth. STP was mainly designed for loop resolution, and original assumptions were that topology changes would be infrequent. STP did not place more emphasis to quick recovery from failures. In data networks, quick recovery is most often not a requirement.

Link Aggregation (LA) IEEE 802.3ad standard is designed to support aggregated links. One of the features Link Aggregation is the support of the possibility of one of the physical link failure in the aggregated link. A Link Aggregation Control Protocol (LACP) is defined "to automatically configure and maintain aggregations among

cooperating systems.” These messages are emitted on a regular, periodic basis. Typically, the period is every second for fast rate and every 30 seconds for slow rate. This means that Link Aggregation also does not support 50 ms recovery. Before the standard was formalized there were several proprietary implementations of link
5 aggregation, most notably Fast EtherChannel product developed by Cisco Systems.

Recently several proprietary technologies have evolved to accomplish the 50 ms second recovery requirement for carrier networks. These technologies can be classified into two main categories: Ethernet based and new non-Ethernet based. In Ethernet
10 based systems most technologies use 20 ms based “Heartbeat” or “Hello” protocol polling to detect link failure along with upper layer software to recover within 50ms. Atrica’s Atrica Resilient Ethernet Access (AREA) framework technology is an example of that. Occam Networks Ethernet Protection Switching (EPS) technology is also an example. Internet Photonics uses interframe gap in Ethernet to support similar
15 functionality.

There are other Ethernet efforts in progress that are also trying to solve the fast recovery problem. Rapid Spanning Tree Protocol (RSTP) IEEE Committee is working on modifications to STP, but currently, there is a requirement of 1 second guaranteed
20 convergence/recovery, not 50 ms. Ethernet First Mile IEEE Committee is also working on modifying Ethernet to support 50ms recovery.

Non Ethernet based technologies being defined include Metro Ethernet Forum, which uses Multi Label Protocol Switching (MPLS) to support protection. Resilient
25 Packet Ring (RPR) technology is being defined by RPR Alliance. RPR is a new protocol that is not compatible with Ethernet protocol, but is designed to support 50 ms recovery in rings.

Most of the above mentioned technologies solve limited functionality for
30 Ethernet. Typically, they support 50 ms protection either in a linear or ring environment, but not both. In addition, they are limited to link failures. They address only a subset of the capabilities as defined by the SDH/SONET APS standard.

SUMMARY OF THE INVENTION

There is a need in the art for APS functionality on an Ethernet network as defined by the SDH/SONET APS standard. The present invention is directed toward
5 further solutions to address this need.

In accordance with one example embodiment of the present invention, Media Access Control (MAC) hardware for supporting MAC Automatic Protection Switching (APS) functionality has a MAC APS Control sublayer and a plurality of MAC sublayers.
10

In accordance with example aspects of the present invention, the plurality of MAC sublayers further includes a Link Aggregation sublayer. The plurality of MAC sublayers can further include at least one MAC Control sublayer. The MAC APS Control sublayer can be located within a MAC. A MAC Control Layer can process
15 optional VLAN tags in control frames. The MAC APS Control sublayer can be implemented in MAC hardware. The MAC APS Control sublayer supports APS for logical links formed by a Link Aggregation sublayer. The MAC APS Control sublayer supports APS for physical links. The MAC APS Control sublayer supports APS for Network Layer paths. A MAC sublayer can be implemented in at least one of an
20 Ethernet Switch device and an Ethernet MAC device. The MAC APS Control sublayer processes an Ethernet MAC control APS frame. The MAC APS Control sublayer maintains an Ethernet MAC Control APS state.

In accordance with further aspects of the present invention, a MAC APS Control
25 Layer generates an interrupt when an APS frame is detected to invoke a MAC Client APS Controller. A MAC APS Control sublayer generates an interrupt when an APS state change is detected to invoke a MAC Client APS Controller. The MAC APS Control sublayer interacts with an Ethernet MAC Client APS Controller.

30 In accordance with another aspect of the present invention, an Ethernet PHY hardware device includes at least one physical sublayer. One of the at least one physical sublayers generates an interrupt when a port/link failure is detected to invoke a MAC Client APS Controller.

In accordance with another aspect of the present invention an Ethernet MAC APS Control Frame for supporting SDH/SONET APS Signalling Protocol includes a standard Ethernet frame Preamble field. Further elements of the Control Frame include
5 a standard Ethernet Start-of-Frame Delimiter field, a standard Ethernet Destination MAC address field, a standard Ethernet Source MAC address field, an optional standard Ethernet VLAN Tag field, a standard Ethernet Type field, a standard Ethernet MAC Control Opcode field, a plurality of standard Ethernet MAC Control Parameters being opcode specific, and a standard Ethernet Frame Check Sequence field.

10

In accordance with further aspects of the present invention the Ethernet MAC Control Opcode further includes an ability to distinguish between a logical link failure, a physical link failure, and a path failure. In addition, the plurality of Ethernet MAC Control Parameters further include a K1 Word field containing an SDH/SONET K1
15 byte, a K2 Word field containing an SDH/SONET K2 byte, a Port ID field, a Slot ID field, a Chassis ID field, a Bridge ID field, a Node ID/IP field, and a Reserved field containing zero or more octets of zero value.

In accordance with further aspects of the present invention, an Ethernet MAC
20 Client includes at least one MAC Client. The MAC Client includes at least one of a network layer protocol and a forwarding function for switches. The MAC Client can also include at least one MAC Control Client APS Controller.

25 In accordance with another aspect of the present invention, a method of providing APS functionality on MAC hardware and PHY hardware includes detecting a failure along a first link on a near end network node. A Physical Layer generates an interrupt when a port/link failure is detected to invoke a MAC Client APS Controller. A switch is made to a second link to correct the failure. The method can execute within 50
30 ms to provide recovery functionality.

In accordance with another aspect of the present invention, a method of providing APS functionality on MAC hardware includes a near end MAC APS Control sublayer receiving a MAC APS Control Frame containing an APS request from a MAC APS Control Frame buffer. The near end MAC APS Control sublayer updates MAC APS state hardware registers to reflect receipt of the APS request. The MAC APS hardware provides maskable interrupts for MAC APS Control Frames received. The near end MAC APS Control sublayer generates interrupts to invoke the APS Controller. The APS Controller processes the APS request. The APS request can include at least one of a switchover request and an APS management request using APS K1/K2 signaling protocol. At least one of manual and automatic switchover APS requests are possible. The method can further include distinguishing between a logical failure, a physical failure, and a path failure. The method can execute within 50 ms to provide recovery functionality.

In accordance with another aspect of the present invention, a method of providing APS functionality on MAC hardware device includes a near end MAC APS Control sublayer receiving APS Controller requests to be transmitted. The MAC APS Control sublayer creates an APS Control frame with requested control parameters. The near end MAC APS Control sublayer transmits the MAC APS Control frame. The APS Controller requests can include at least one of a switchover request and an APS management request using APS K1/K2 signaling protocol. The APS Controller requests can also include at least one of a manual and automatic switchover APS request. The method can execute within 50 ms to provide recovery functionality.

In accordance with another aspect of the present invention, a method of providing APS functionality on an Ethernet protocol network includes experiencing a failure along a first port/link. An interrupt is generated. The interrupt is forwarded to an APS controller. The APS controller initiates a switch from the first port/link to a second port/link. The method can execute within 50 ms to provide recovery functionality.

In accordance with another aspect of the present invention, a method of providing APS functionality on an Ethernet protocol network includes receiving an APS Control frame with an APS request. The APS Control frame APS request is processed.

An interrupt is generated. The interrupt is forwarded to an APS Controller, the APS Controller processing the APS request received. The method can execute within 50 ms to provide recovery functionality. In addition, the method can provide support for standard SDH/SONET APS functionality for linear, ring, and mesh topologies for
5 Ethernet protocol networks using SDH/SONET K1/K2 bytes for SDH/SONET APS Signaling protocol.

In accordance with another aspect of the present invention, a method of providing APS functionality on an Ethernet protocol network includes receiving an APS
10 request from an APS Controller Client. An APS Control Frame is created with the APS request. The APS Control Frame is transmitted. The method can execute within 50 ms to provide recovery functionality. The method can further provide support for standard SDH/SONET APS functionality for linear, ring, and mesh topologies for Ethernet
15 protocol networks using SDH/SONET K1/K2 bytes for SDH/SONET APS Signaling protocol.

BRIEF DESCRIPTION OF THE DRAWINGS

The aforementioned features and advantages, and other features and aspects of
20 the present invention, will become better understood with regard to the following description and accompanying drawings, wherein:

FIGS. 1A and 1B illustrate the switching of a path using APS according to one aspect of the present invention;

FIG. 2 is a diagrammatic illustration of an Ethernet MAC APS Control
25 architecture according to one embodiment of the present invention;

FIG. 3 is a diagrammatic illustration of an Ethernet MAC APS Control sublayer internal architecture according to one aspect of the present invention;

FIG. 4 is a diagrammatic illustration of an Ethernet MAC APS Control Frame format according to one aspect of the present invention;

FIG. 5 is a diagrammatic illustration of Ethernet MAC APS Operation according
30 to one embodiment of the present invention; and

FIG. 6 is a diagrammatic illustration of Ethernet MAC APS Control Flow according to one embodiment of the present invention.

DETAILED DESCRIPTION

5 Illustrative embodiments of the present invention relate to the implementation of standard SDH/SONET APS functionality within an Ethernet architecture. In order to support SDH/SONET APS functionality in the Ethernet architecture, the present invention extends the Ethernet MAC Control Sublayer. The Ethernet MAC Control sublayer is a sublayer of the data link layer (Layer 2, described later herein). The MAC Control sublayer resides between the MAC (the Media Access Control, which is an
10 entity or algorithm utilized in negotiating access to a shared or dedicated communications channel) and a client of that MAC (where the client is typically a network layer protocol or a relay function implemented by bridges or switches). The clients of the MAC can use the MAC Control sublayer to control the operation of the Ethernet MAC. The implementation of MAC Control sublayer is optional under
15 Ethernet standards.

Aspects of the present invention include an Ethernet MAC APS Control Protocol, which can be used to support Ethernet MAC APS. The Ethernet MAC APS Control Protocol extends the MAC Control sublayer to make use of Ethernet MAC
20 multicast or unicast addresses, and MAC Control opcodes to support the APS function. The MAC APS function implements SDH/SONET APS on full duplex Ethernet links. The MAC APS frame contains the K1/K2 bytes as described by the SDH/SONET APS standards. The MAC APS also operates in a same manner to the known SDH/SONET APS. An APS Controller can be the client for the Ethernet MAC APS Control sublayer.
25 The APS Controller uses the Ethernet MAC APS infrastructure provided in accordance with aspects of the present invention to implement standard APS functionality.

FIGS. 1 through 6, wherein like parts are designated by like reference numerals throughout, illustrate example embodiments of methods for implementing SDH/SONET
30 APS in an Ethernet environment, according to the present invention. Although the present invention will be described with reference to the example embodiments illustrated in the figures, it should be understood that many alternative forms can embody the present invention. One of ordinary skill in the art will additionally

appreciate different ways to alter the parameters of the embodiments disclosed, such as the size, shape, or type of elements or materials, in a manner still in keeping with the spirit and scope of the present invention.

5 In order to appreciate operation of the illustrative embodiments described herein, it is helpful to understand the Open Systems Interconnect (OSI) network hierarchy, which views a network as being composed of several hierarchical layers. In the hierarchy, Layer 1 is the physical layer containing elements that perform the transmission of signals within the network. Layer 2 is the data link layer, which
10 provides services that allow direct communication between devices across the underlying physical channel of Layer 1. Layer 3 is the network layer, which is responsible for station-to-station data delivery over multiple data links. The network layer is responsible for the routing of packets across the network. Layer 4 is the transport layer, which provides an error-free, sequenced, guaranteed delivery, message
15 service that allows process to process communication between stations on a network. Layer 5 is the session layer, which deals with the establishment of communications between applications. This layer is useful for security applications. Layer 6 is the presentation layer, which enables the sharing of data between networked systems using different methods of local data representation. Finally, Layer 7 is the application layer.
20 This layer provides generic application functions, such as email, file transfer capability, and the like.

 In SDH/SONET, APS provides port/line protection between nodes at a physical layer (i.e., Layer 1). Several topologies can support APS (i.e., ring, linear, or mesh
25 topologies) and several levels of protection are possible (i.e., 1+1, N+1, and N:1). For the purpose of the following description, APS is illustrated using 1+1 APS configuration between nodes. N+1 and N:1 APS configurations can be supported by the same architecture. The APS "1 + 1" architecture generally arranges two lines or paths, with information propagating down each line or path at the same time. The connection can
30 be bi-directional or unidirectional. In addition, the "1 + 1" architecture provides for a protect interface, or circuit, paired with each working interface, or circuit. Often, the protect and working circuits interface with an add/drop multiplexer, which sends the same traffic load to the working and protect circuits.

Within the protect circuit, information indicating the current status of the APS connection travels through the circuit continuously and conveys any requests for action. This information can be used to synchronize the working and protect circuits.

5

The present invention can support linear and ring protection, ring and mesh topology, and provides physical links, logical links, and path protection. However, for purposes of clarity in describing the invention, the description contained herein utilizes a linear APS configuration. One of ordinary skill in the art will appreciate that ring and mesh topology protection can also be implemented in accordance with aspects of the present invention. Path level protection can also be supported by the architecture of the present invention by providing the appropriate source and destination Ethernet MAC addresses.

15 **FIGS. 1A and 1B** illustrate an APS configuration between two nodes. There are two network elements or nodes, a first node 102 and a second node 103. The first node 102 has a first fiber pair extending along a first link 104 and the second node 103 has a second fiber pair extending along a second link 105. The first link 104 and the second link 105 connect the first and second nodes 102 and 103 respectively at a first APS port pair 110 and a second APS port pair 112. Inside each node 102 and 103, there are connections 106, 107, 108, and 109 between ingress and egress ports of the nodes 102 and 103. The solid line arrows represent active traffic, while the dashed line arrows represent standby traffic. For simplicity, **FIGS. 1A and 1B** only show traffic flowing in one direction. In the case of bi-directional architecture, the other direction has the same traffic pattern in the opposite direction from the arrows illustrated. **FIGS. 1A and 1B** show before and after states of the APS traffic (i.e., before a line interruption and after a fiber has been cut at fiber cut 111, causing a line interruption).

30 In APS, as shown in **FIGS. 1A and 1B**, one link (104 or 105) is protected by another link (104 or 105) to anticipate and address various kinds of failures. These failures can include equipment failures such as node failures, card failures, and port failures, or link failures, such as a cable/fiber cut. These errors are commonly known as

hard failures. A second category of errors called soft failures includes instances when significant bit error rates occur on a link.

5 The first link 104 supports the working port/link, which extends between the first
node 102 and the second node 103. The second link 105 supports the protect port/link,
which also extends between the first node 102 and the second node 103. The
determination of which link 104 or 105 is active as the working port/link and the protect
port/link depends upon the state of the APS controller on nodes 102 and 103. In FIG.
1A, the first link 104 is the active working port/link. The frames of data are transmitted
10 to both a first working connection 106 and a first protect connection 107. The
transmission of the frames is known as bridging. The frames of data propagate along the
working port/link of the first link 104 and the protect port/link of the second link 105.
The frames of data then transition through the second node 103 along an active second
working connection 108, but do not propagate through a standby second protect
15 connection 109. A selector can select which path is used for receiving frames. Bridging
and Selector can be implemented using hardware which supports dual casting, such as Y
connectors for electrical interfaces, 2x2 cross connect switch chips for electrical or
optical interfaces. Serial bus architecture can also be used.

20 When a failure occurs on the working port/link along the first link 104, for
example due to the fiber cut 111 of FIG. 1B, a receiver (not shown) in the second node
103 detects the link failure in hardware and causes an interrupt for the MAC APS Client,
which in turn initiates a switchover to the protect port/link of the second link 105. The
second working connection 108 enters a standby condition, and the second protect
25 connection 109 becomes active. The frames of data can then continue between the first
node 102 and the second node 103, by propagation along the protect port/link of the
second link 105 from the first protect connection 107 to the second protect connection
109 of the second node 103.

30 FIG. 2 shows an Ethernet MAC APS Control Architecture. The MAC Client or
Higher Layers 201 can be, e.g., a network layer protocol, such as IP, or a forwarding
function for switches. In the illustrated embodiment, the MAC Client 201 also
implements the APS Controller functionality. The APS Controller in the MAC Client

201 manages the state of the APS and reacts to various errors or commands to switchover.

A MAC APS Control sublayer 202 supports APS for logical links formed by
5 Link Aggregation sublayer 203. The Link Aggregation sublayer 203 allows a plurality of physical links to be aggregated into one aggregated link. An aggregated link is one form of a logical link. In addition, MAC APS Control sublayers 204A, 204B, and 204C support APS for physical links. Standard MAC Control sublayers 205A, 205B, and 205C, support all the currently defined MAC control frames, e.g., PAUSE frames, which
10 prevent switches from unnecessarily discarding data frames due to input buffer overload. All of the MAC Control sublayers 205A, 205B, and 205C are optional. A standard MAC sublayer 206A, 206B, and 206C, controls access to media. A standard PHY sublayer 207A, 207B, and 207C, implements physical layer signals for transmission media.

15

For SDH/SONET APS to support bi-directional switchover, preemption, and several other APS features, SDH/SONET APS requires support of APS K1/K2 signaling protocol. K1/K2 signaling protocol relates to the actual bytes used in SDH/SONET signaling. More specifically, the K1 byte and the K2 byte in the SDH/SONET
20 architecture are used for protection signaling between line terminating entities for bi-directional automatic protection switching, and for detecting alarm indication signals (AIS-L) and remote defect indication signals (RDI).

The MAC Control Sublayer extends to provide APS functionality by enabling
25 Ethernet to support K1/K2 signaling protocol. The MAC Control sublayer also extends to provide optional support of VLAN tags for MAC control frames. Aspects of the present invention introduce the Ethernet MAC APS Control sublayers 202, 204A, 204B, and 204C, as shown in FIG. 2. The Ethernet MAC APS Control sublayers 202, 204A, 204B, and 204C, include an Ethernet MAC Control APS frame, which contains the
30 K1/K2 bytes in accordance with definitions of the K1/K2 bytes in the known SDH/SONET standards. The MAC APS sublayer 202 can support MAC APS at Link Aggregation sublayer 203 (logical network interface layer) and/or physical network interface layer 204A, 204B, and 204C. It should be noted that the same physical

hardware can be used to process the APS signaling protocol for the MAC APS sublayer 202, 204A, 204B, and 203C.

APS in Ethernet can be implemented at the physical layer (Layer 1) like
5 SDH/SONET in the framing process by modifying the framing or using interframe gaps for APS signaling protocol, or it can be implemented in the Ethernet MAC Control sublayer like the PAUSE function or Link Aggregation function known in the art. Modifying Ethernet framing to implement APS would not be backward compatible and would be difficult to standardize through the end users and industry. Also, modifying
10 Ethernet framing would be only a partial solution because Ethernet frames are only transmitted when there is data to be sent. If there is no data to be sent, there is no Ethernet frame to carry APS signals. Contrarily, in SDH/SONET, frames are continuously generated (data or idle). Therefore, implementing APS at the Ethernet MAC Control sublayer allows the APS signaling to be event and packet driven in a
15 natural Ethernet manner.

FIG. 3 shows the Ethernet MAC APS Control sublayer internal architecture. Again, MAC Client or Higher Layers 306 can be a network layer protocol, such as IP, or a forwarding function for switches. In the illustrated embodiment, the MAC client 306
20 again implements MAC APS Controller 305 functionality. The MAC APS Controller 305 manages the state of the APS and reacts to various errors or commands to switchover.

A MAC APS Control sublayer 301 supports APS for logical links formed by a
25 Link Aggregation sublayer 304. The Link Aggregation sublayer 304 allows many physical links to be aggregated into one logical link. MAC APS Control sublayers 302A, 302B, and 302C support APS for physical links. Standard MAC Control sublayers 307A, 307B, and 307C, support all of the currently defined MAC control frames, e.g., PAUSE frames. All the MAC Control sublayers are optional. Standard
30 MAC sublayers 308A, 308B, and 308C control access to the media. Standard PHY sublayers 303A, 303B, and 303C, implement physical layer signals for transmission media.

The MAC APS Control sublayer 301 includes a MAC APS Control operation 301A, which manages a MAC APS state 301C based on MAC APS Control frames received from MAC APS Control Frame buffers 301B. The MAC APS Control operation 301A also generates an interrupt along line 314 when there is a change in APS
5 K1/K2 state. The MAC APS Controller 305 can access the MAC APS State 301C. The MAC APS State 301C contains various interrupt status registers, K1/K2 byte state registers, APS opcode received, and other APS related information. The MAC APS Controller 305 can also transmit MAC APS Control Frames from MAC APS Control Frame buffers 301D and built by the MAC APS Controller 305 by setting up registers in
10 the MAC APS State 301C.

As shown in **FIG. 3**, the Ethernet MAC Architecture is modified to support MAC APS Control sublayers 301, 302A, 302B, and 302C. The Ethernet MAC architecture supports APS processing in the MAC APS Control Operation 301A of
15 MAC APS Control Frame buffers 301B and 301D, maintains MAC APS State 301C, and generates interrupts along interrupt line 314 when a MAC APS Control Frame has been received and/or if there has been a change in APS state at the MAC APS State 301C.

20 The Ethernet hardware supports interrupts from Physical Layers 303A, 303B, and 303C from hard failures due port failures or link failures. Similarly, the Ethernet hardware can be extended to support soft failures, such as error rates greater than predetermined configured thresholds. The Ethernet hardware provides additional registers at the MAC APS State 301C for higher Layers and the MAC APS Controller
25 305 to access, via the access line 313, the current state of APS, such as K1/K2 bytes, and the like.

The dedicated MAC APS Control Frame buffers 301B and 301D receive and transmit, respectively, the MAC APS Control Frames. The existence of the MAC APS
30 Control Frame buffers 301B and 301D prevents head of queue blocking of the control frame so that a link switchover can occur within 50 ms, in accordance with the SDH/SONET standard. The Ethernet hardware can additionally provide separate control and maskable status registers for APS functionality in the MAC APS State 301C. The

MAC APS Control Frames (transmit) from the MAC APS Control Frame buffer 301D can be accessed by the MAC APS controller 305 using hardware path 312. The access can be implemented as registers or via direct access to the MAC APS Control Frame buffer 301D. The received data frames directly pass along the Client Frames receive path 310 to the MAC client 306. The MAC Client 306 transmit data frames path 311 transmits the MAC Client 306 data frames directly the MAC hardware.

The MAC APS Control sublayer as shown in FIG. 3 can be placed below the Link Aggregation sublayer 304 to protect individual physical links of the MAC APS Control sublayer 302A, 302B, and 302C. Alternatively, the MAC APS Control sublayer can be positioned above the Link Aggregation sublayer 304 to protect logical links. The same architecture supports both cases. Different MAC Control opcodes are utilized to distinguish between each of the different cases.

When a failure is detected locally on a near end node (the node closer to the source of the transmission) the Ethernet PHY layer 303A, 303B, and 303C generates an interrupt along interrupt line 314, which invokes the MAC APS Controller 305 and in turn causes a switchover.

When a failure is detected remotely by a far end node (a node closer to the destination of the transmission) and requests a switchover, it sends a MAC APS Control Frame from the MAC APS Control Frame buffer 301D with the appropriate APS command and information as needed by the K1/K2 bytes. When the MAC APS Control Frame is received at the near end node, the near end MAC Control sublayer 301A, 302A receives the MAC APS Control Frame from the MAC APS Control Frame buffer 301B and updates the MAC APS hardware state/registers at the MAC APS State 301C to reflect the new request. The hardware provides maskable interrupts for MAC APS Control Frames received and if there is change in MAC APS State bytes at the MAC APS State 301C. When such an interrupt occurs it again invokes the MAC APS Controller 305, which executes the APS switchover.

In the transmit direction, the hardware can be implemented such that there are K1/K2 registers at the MAC APS State 301C and when the APS Controller 305 needs to transmit the MAC APS Control Frame it writes to the MAC APS Control Frame buffer 301D via the hardware path 312 of the control registers.

5

There are several embodiments for implementation of aspects of the present invention in the MAC layer. To distinguish which MAC APS context exists in any one instance, logical versus physical, one of several processes below occurs. One implementation is to use the MAC Control Frames, multicast destination address 01-80-10 C2-00-00-03, type 0x8808 and three opcodes for Physical APS, Logical APS frames, and Path APS frames. A second implementation can make use of the different frame types instead of opcodes. Alternatively, an embodiment can implement MAC APS as one class of Slow Protocols, as defined in the Ethernet Standards, a class of protocols wherein they never emit more than a specified maximum number of frames per time 15 period. Still, another embodiment includes implementation using vendor specific; multicast destination addresses. One of ordinary skill in the art will appreciate that different combinations of the above embodiments, as well as others not specified, can also be used. One example embodiment, detailed below, shows implementation of the invention with the first option.

20

FIG. 4 shows an example Ethernet MAC APS Control Frame 400 format. All Ethernet frames start with seven bytes of Preamble 401, each containing the value 0x55. A Start of Frame Delimiter (SFD) 402 contains the value 0xD5. A destination address 403 contains the unique multicast address reserved for MAC APS operations: 25 01-80-C2-00-00-03. This would require registering with the 802 Standards Committee. The destination address 403 can also be the unicast MAC address of the destination port. The destination port would be configured through an external mechanism, such as the system software. The destination address 403 requires six bytes. A source address 404, also requiring six bytes, contains the unicast address of the source interface sending the 30 MAC APS frame. A VLAN tag field 410 is optional and contains standard VLAN Protocol ID 0x8100 in the first two bytes of the field and the second two bytes contain the VLAN Identifier, priority, and Tag Control Info bit. These are defined by the 802.1Q/1p standards. A type field 405 contains the reserved value 0x8808 used for all

MAC Control Frames, and requires two bytes. A MAC APS Control Opcode 406 for MAC APS physical link level is 0x0002. The MAC APS Control Opcode 406 for MAC APS logical link level is 0x0003. The MAC APS Control Opcode 406 for MAC APS path level is 0x0004. This would require registering with the 802 Standards Committee.

5 In all instances, the MAC APS Control Opcode 406 requires two bytes. A MAC APS Control Parameters field 407 takes two parameters called K1 Word 407A and K2 Word 407B. These are 4-byte unsigned integer values containing the K1 and K2 bytes of standard SDH/SONET APS. The use of word length allows for growth in the K1 byte and allows more than 16 Station IDs in the K2 byte. Within the MAC APS Control

10 Parameters, additional optional fields exist for Port ID 407C, Slot ID 407D, Chassis ID 407E, Bridge ID 407F and Node ID/IP 407G. These fields can be used for fault isolation in case of path APS protection, and require the bytes illustrated in the figure. External software can use these fields to generate alarms or report the status of where the failure occurred. A Reserved field 408 is maintained for future extensions and is set to

15 all zeros. A Frame Check Sequence (FCS) field 409 is a checksum computed on the contents of the frame from the Destination Address 403 through to the end of the Reserved field 408 inclusively.

FIG. 5 illustrates one example implementation of the Ethernet MAC APS

20 Operation. MAC Client or Higher Layers 501 can be a network layer protocol, such as IP, or a forwarding function for switches. In the illustrated embodiment, the MAC client 501 also implements MAC APS Controller 501A functionality. The MAC APS Controller 501A manages the state of the APS and reacts to various errors or commands to switchover. MAC APS Control sublayers 503 and 514 support APS for logical links

25 formed by Link Aggregation sublayers 504 and 515. Link Aggregation sublayers 504 and 515 allow many physical links to be aggregated into one logical link. MAC APS Control sublayers 505, 509, 516, and 520 support MAC APS for physical links. In a next layer is a standard MAC Control sublayer 506, 510, 517, and 521, which supports all the currently defined MAC APS Control Frames, e.g., PAUSE frames. All the MAC

30 Control sublayers are optional. After the MAC Control sublayer is a standard MAC sublayer 507, 511, 518, and 522, that controls access to the media. Next is a standard PHY sublayer 508, 512, 519, and 523, that implements the physical layer signals for the transmission media.

A logical level working link 502 transmit 502A and receive 502B are shown as line arrows representing active traffic. A logical level protect link 513 transmit 513A is shown active and receive 513B is shown in standby mode. Similarly, each of the
5 physical links are illustrated as physical working links 523 and 525, corresponding active transmit links 523A and 525A, and corresponding active receive links 523B and 525B. Physical protect links 524 and 526 are also shown, with corresponding active transmit links 524A and 526A, and corresponding standby receive links 524B and 526B.

10 The APS operation is as shown in **FIG. 5**, which shows the case of logical (link aggregation) APS links as working links 502 and protect links 513. **FIG. 5** also shows the case of physical APS links as working links 523 and 525, and protect links 524 and 526. The APS Controller 501A follows the same state machines and implements the same commands as standardized in the SDH/SONET standards referenced.

15 The APS operation control flow is illustrated in **FIG. 6** while concomitantly referring to **FIGS. 4** and **5** in describing the operation of the present invention. First a determination is made whether there is a failure detected (step 610). If no failure has been detected, no action is taken (step 612). If there is a local failure detected, such as
20 link failure or bit errors crossing a threshold on working link 523, the Ethernet MAC, 507 or PHY 508 layer raises an interrupt for the MAC APS Controller 501A (step 620). The MAC APS Controller 501A is then invoked (step 630). The MAC APS Controller 501A causes all the failed physical links to switch over from the working link 523 to the protect 524 link (step 640). This can be done for all physical links associated with the
25 logical working link 502. A determination is again made as to whether there is still a failure (step 650). If there is no more failure, no further action is taken (step 652). If the logical link continues to receive errors, then the MAC APS controller can cause the logical link switchover from logical working link 502 to logical protect link 513 (step 660).

30 When a remote or far end node wants to request a switchover, it sends an Ethernet MAC APS frame with the appropriate APS command in the K1/K2 bytes (steps 642, 662, and 682). When a near end node receives the frame, the appropriate

switchover is executed. If the opcode is 0x0002, then the physical link switches over (executing step 640). If the opcode is 0x0003, then the logical link switches over (executing step 660). If the opcode is 0x0004, then the path switches over (executing step 680).

5

The above-described functionality allows for path level protection in ring topology as defined by the SDH/SONET standards. In order to support path level APS in mesh topology, the method can continue as follows.

10 Using the unicast MAC addresses of the end points for the destination address 403 and the path level APS opcode 0x0003 in the opcode field 406 in FIG. 4, path level APS can be supported. The unicast MAC address can be that of the working or protect port, or that of the port being protected by APS. When an APS path is configured, external software then provides the path to the APS controller. The information can also
15 be stored in the MAC APS state registers and used to create the path APS control frame. In the example case of path level APS support, the APS frame is sent to the two end points of the circuit, where the MAC hardware receives the frame with the path APS opcode specified in the MAC APS Control Frame. In order for this to occur, a higher level application can already have caused the path to the end points to be learned and not
20 aged until the APS path configuration is removed. Alternatively, the higher level application can continue to rely on learning, although learning may not be able to achieve sub 50ms restoration times. To provide priority to the APS frame 400, the VLAN field 410 can be set up with appropriate priority.

25 To continue with the method as illustrated in FIG. 6, a determination is again made as to whether a failure still exists (step 670). If no failure exists, no action is taken (step 672). If a failure does exist, the APS frame 400 goes through the generic APS control mechanism and invokes the MAC APS Controller 501A, which detects that this was received as a path level request (step 682) and executes the path level switchover
30 (step 680). The MAC APS Controller 501A then invokes the higher application, which updates its path state and additional post switchover processing. Post switchover processing could include configurations for finding alternative protect paths, or other functionality.

It should be noted that if the VLAN tag 410 is used for path level protection, then by definition, all the stations/nodes along the path need to be VLAN aware.

Alternatively, the edge station/node can strip off the VLAN tag. Configuration must
5 indicate that this is a VLAN edge station/node.

The APS frame 400 can additionally be constructed to include more information such as the Port ID 407C, Slot ID 407D, Bridge ID 407F, and node ID/IP 407G. These can be used for fault isolation.

To increase reaction speed, there can be two MAC APS Control Frame buffers, one for link level protection and one for path level protection. Because these two MAC APS Control Frames can be significantly different, they should not be constructed at the moment of failure. The logical and physical link frame are very similar so there might
10 not be need for additional Frame buffer between the logical and physical link frames. The trade off for the implementation balance cost and efficiency.

In the instance where all three protections (i.e., physical level protection, link level protection, and path level protection) are in operation, the physical link protection
20 is first to react (executing step 640). If the failure continues to persist (as determined at step 650), the logical link protection reacts (executing step 660). If the failure continues to persist (as determined at step 670), then path level protection is triggered, and the end stations execute a path level protection switchover (executing step 680). If a subsequent determination finds that there is no longer a failure, no further action is taken (step 692).
25 However, if a failure is still detected, then an alarm is raised by external software for circuit failure (step 695), and concluding the method of operation as it relates to the present invention.

The present invention maintains several advantages. Because it does not change
30 the Ethernet framing it is backward compatible and conforms to the Ethernet Standards. Because it modifies the optional Ethernet MAC Control sublayer, vendors can choose whether to implement the functionality of the present invention. The elegant design and configuration of the invention makes implementation in the MAC hardware relatively

easy. The method of the present invention is event driven, thus APS frames are only sent when required. Therefore, the apparatus of the present invention maintains low overhead and consumes very little bandwidth. Because it reuses standard SDH/SONET APS definitions and process, the implementation of the present invention is fully
5 compliant with the SDH/SONET standard protocol. Because it supports pre-configured frames and is not head of queue blocking, the present invention can be implemented to support switchover in 50 ms or less. In addition, the present invention makes use of standard SDH/SONET APS definitions and process, therefore it can support protection for linear, mesh, and ring topologies, physical level, logical level, and path level
10 protection, and 1+1, N+1, and N:1 protection.

Numerous modifications and alternative embodiments of the present invention will be apparent to those skilled in the art in view of the foregoing description. Accordingly, this description is to be construed as illustrative only and is for the purpose
15 of teaching those skilled in the art the best mode for carrying out the present invention. Details of the structure may vary substantially without departing from the spirit of the invention, and exclusive use of all modifications that come within the scope of the appended claims is reserved. It is intended that the present invention be limited only to the extent required by the appended claims and the applicable rules of law.

CLAIMS

What is claimed is:

- 5 1. A Media Access Control (MAC) hardware device for supporting MAC Automatic Protection Switching (APS) functionality, comprising:
a MAC APS Control sublayer for processing APS frames and maintaining APS state; and
a plurality of MAC sublayers for carrying out MAC operations.
- 10 2. The MAC hardware of claim 1, wherein the plurality of MAC sublayers further comprises a Link Aggregation sublayer.
- 15 3. The MAC hardware of claim 1, the plurality of MAC sublayers further comprises at least one MAC Control sublayer for generating control frames.
- 20 4. The MAC hardware of claim 1, wherein the MAC APS Control sublayer is located within a MAC.
- 25 5. The MAC hardware of claim 1, wherein a MAC Control Layer processes optional VLAN tags in control frames.
6. The MAC hardware of claim 1, wherein the MAC APS Control sublayer is implemented in the MAC hardware.
7. The MAC hardware of claim 1, wherein the MAC APS Control sublayer supports APS for logical links formed by a Link Aggregation sublayer.
8. The MAC hardware of claim 1, wherein the MAC APS Control sublayer supports APS for physical links.
- 30 9. The MAC hardware of claim 1, wherein the MAC APS Control sublayer supports APS for Network Layer paths.

10. The MAC hardware of claim 9, wherein the MAC sublayer can be implemented in at least one of an Ethernet Switch device and an Ethernet MAC device.
- 5 11. The MAC hardware of claim 1, wherein the MAC APS Control sublayer processes an Ethernet MAC control APS frame.
12. The MAC hardware of claim 1, wherein the MAC APS Control sublayer maintains an Ethernet MAC Control APS state.
- 10 13. The MAC hardware of claim 1, wherein a MAC APS Control Layer generates an interrupt when an APS frame is detected to invoke a MAC Client APS Controller.
14. The MAC hardware of claim 1, wherein a MAC APS Control Layer generates an interrupt when an APS state change is detected to invoke a MAC Client APS Controller.
- 15 15. The MAC hardware of claim 1, wherein the MAC APS Control sublayer interacts with an Ethernet MAC Client APS Controller.
- 20 16. An Ethernet PHY hardware device, comprising:
a decoder; and
at least one physical sublayer, wherein one of the at least one physical sublayers generates an interrupt when a port/link failure is detected to invoke a MAC Client APS Controller.
- 25 17. In a MAC APS hardware device, an Ethernet MAC APS Control Frame for supporting SDH/SONET APS Signalling Protocol, comprising:
a standard Ethernet frame Preamble field;
a standard Ethernet Start-of-Frame Delimiter field;
30 a standard Ethernet Destination MAC address field;
a standard Ethernet Source MAC address field;
an optional standard Ethernet VLAN Tag field;
a standard Ethernet Type field;

a standard Ethernet MAC Control Opcode field;
a plurality of standard Ethernet MAC Control Parameters being opcode specific;
and
a standard Ethernet Frame Check Sequence field.

5

18. The Ethernet MAC APS Control Frame of claim 17, wherein the Ethernet MAC Control Opcode further comprises an ability to distinguish between a logical link failure, a physical link failure, and a path failure.

10 19. The Ethernet MAC APS Control Frame of claim 17, wherein the plurality of Ethernet MAC Control Parameters are further comprised of:

a K1 Word field containing an SDH/SONET K1 byte;
a K2 Word field containing an SDH/SONET K2 byte;
a Port ID field;
15 a Slot ID field;
a Chassis ID field;
a Bridge ID field;
a Node ID/IP field; and
a Reserved field containing zero or more octets of zero value.

20

20. An Ethernet MAC Client, comprising:

at least one of a network layer protocol and a forwarding function for switches;
and
at least one MAC Control Client APS Controller.

25

21. A method of providing APS functionality on a MAC hardware and PHY hardware, comprising:

detecting a failure along a first link on a near end network node;
a Physical Layer generating an interrupt when a port/link failure is detected to
30 invoke a MAC Client APS Controller; and
switching to a second link to correct the failure.

22. The method of claim 21, wherein the method executes within 50 ms to provide recovery functionality on an Ethernet protocol network.
23. A method of providing APS functionality on a MAC hardware, comprising:
- 5 a near end MAC APS Control sublayer receiving a MAC APS Control Frame containing an APS request from a MAC APS Control Frame buffer;
- the near end MAC APS Control sublayer updating MAC APS state hardware registers to reflect receipt of the APS request;
- the MAC APS hardware providing maskable interrupts for MAC APS Control
- 10 Frames received;
- the near end MAC APS Control sublayer generating interrupts to invoke the APS Controller; and
- the APS Controller processing the APS request.
- 15 24. The method of claim 23, wherein the APS request comprises at least one of a switchover request and an APS management request using APS K1/K2 signaling protocol.
- 20 25. The method of claim 23, further comprising at least one of manual and automatic switchover APS requests.
26. The method of claim 23, further comprising distinguishing between a logical failure, a physical failure, and a path failure.
- 25 27. The method of claim 23, wherein the method executes within 50 ms to provide recovery functionality.
28. A method of providing APS functionality on a MAC hardware device, comprising:
- 30 a near end MAC APS Control sublayer receiving APS Controller requests to be transmitted;
- the MAC APS Control sublayer creating an APS Control frame with requested control parameters; and

the near end MAC APS Control sublayer transmitting the MAC APS Control frame.

29. The method of claim 28, wherein the APS Controller requests comprise at least one
5 of a switchover request and an APS management request using APS K1/K2 signaling protocol.

30. The method of claim 28, wherein the APS Controller requests comprise at least one
10 of a manual and automatic switchover APS request.

31. The method of claim 28, wherein the method executes within 50 ms to provide recovery functionality.

32. A method of providing APS functionality on an Ethernet protocol network,
15 comprising the steps of:
 experiencing a failure along a first port/link;
 generating an interrupt; and
 forwarding the interrupt to an APS controller, the APS controller initiating a switch from the first port/link to a second port/link.

20 33. The method of claim 32, wherein the method executes within 50 ms to provide recovery functionality.

34. A method of providing APS functionality on an Ethernet protocol network,
25 comprising the steps of:
 receiving an APS Control frame with an APS request;
 processing the APS Control frame APS request;
 generating an interrupt; and
 forwarding the interrupt to an APS Controller, the APS Controller processing the
30 APS request received.

35. The method of claim 34, wherein the method executes within 50 ms to provide recovery functionality.

36. The method of claim 34, wherein the method provides support for standard SDH/SONET APS functionality for linear, ring, and mesh topologies for Ethernet protocol networks using SDH/SONET K1/K2 bytes for SDH/SONET APS Signaling
5 protocol.

37. A method of providing APS functionality on an Ethernet protocol network,
comprising of steps of:
receiving an APS request from an APS Controller Client;
10 creating an APS Control Frame with the APS request; and
transmitting the APS Control Frame.

38. The method of claim 37, wherein the method executes within 50 ms to provide
recovery functionality.
15

39. The method of claim 37, wherein the method provides support for standard SDH/SONET APS functionality for linear, ring, and mesh topologies for Ethernet protocol networks using SDH/SONET K1/K2 bytes for SDH/SONET APS Signaling
protocol.
20

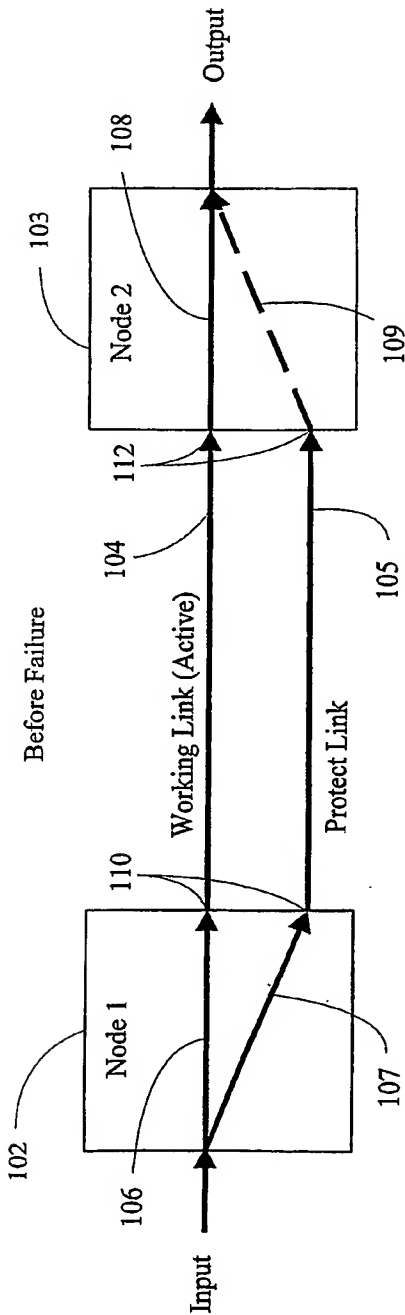


FIG. 1A

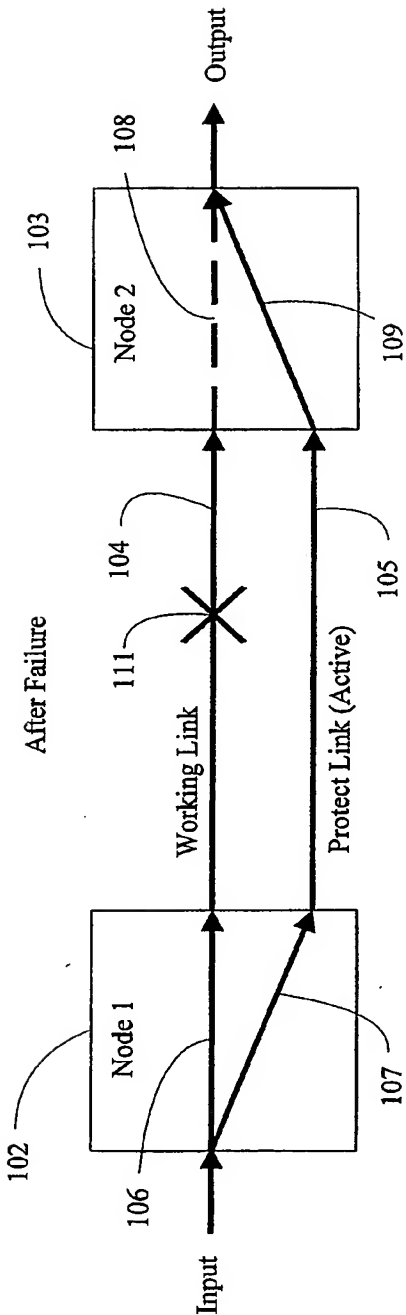


FIG. 1B

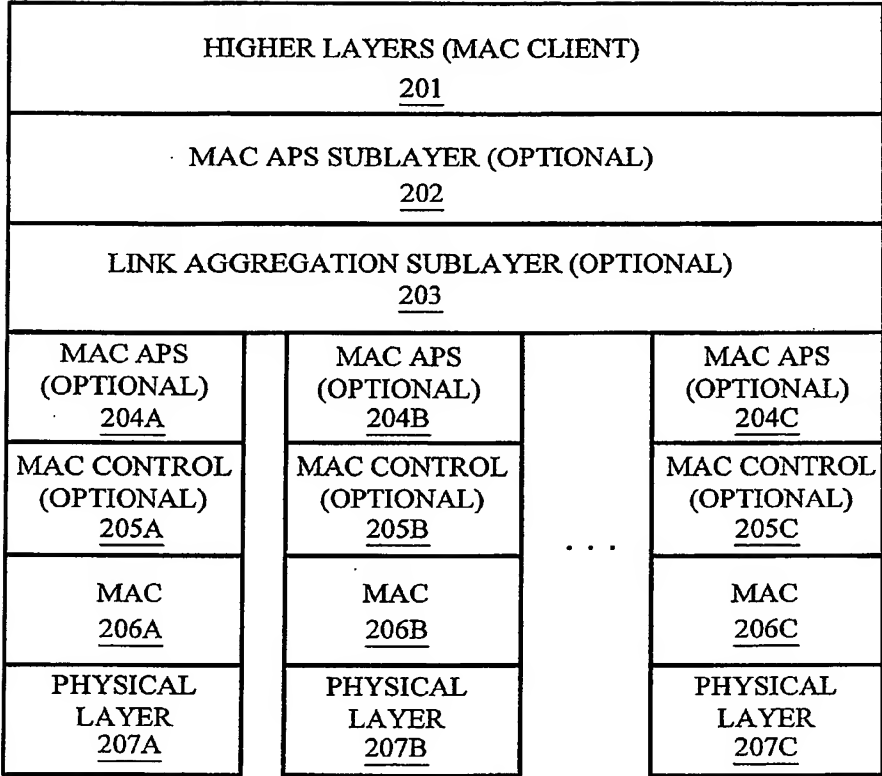


FIG. 2

3/6

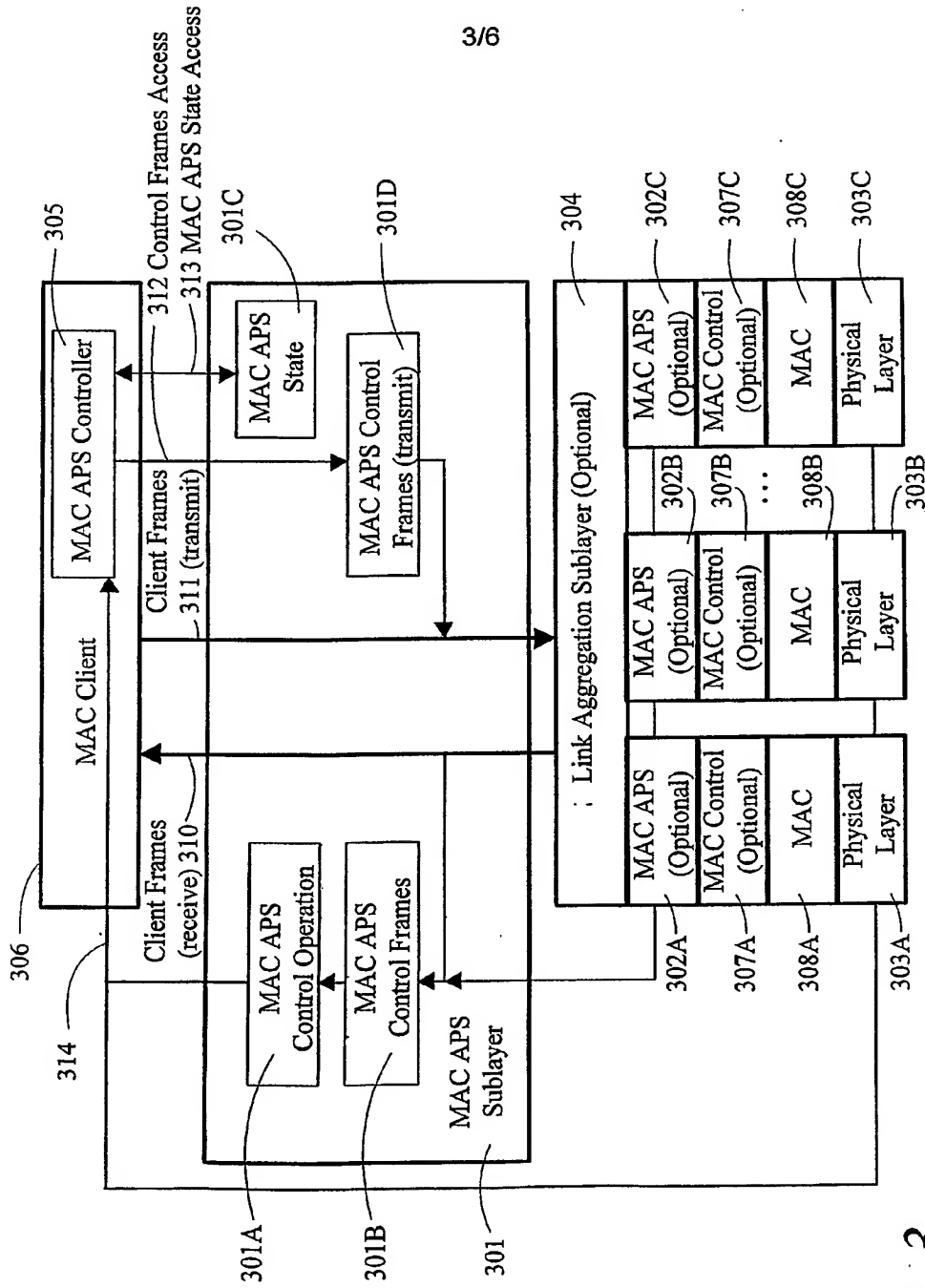
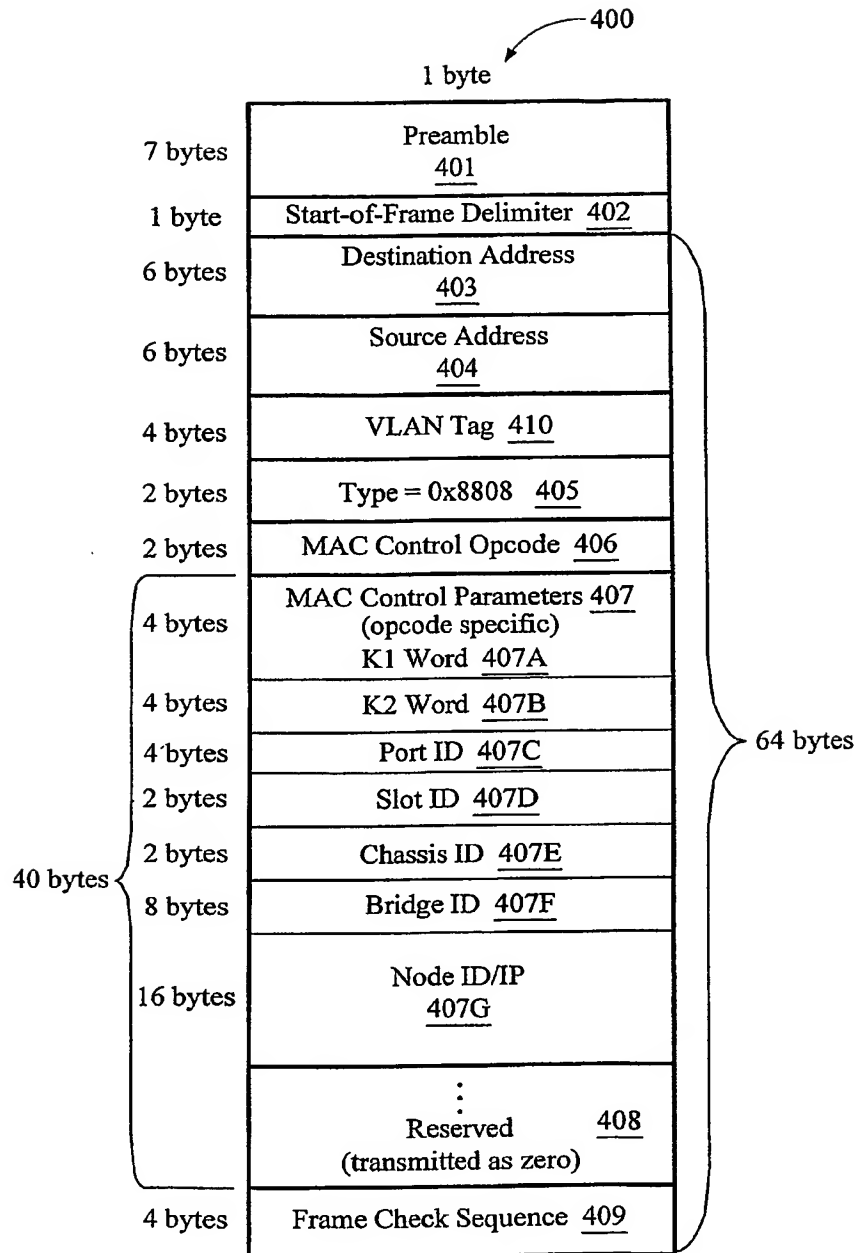


FIG. 3

4/6

*FIG. 4*

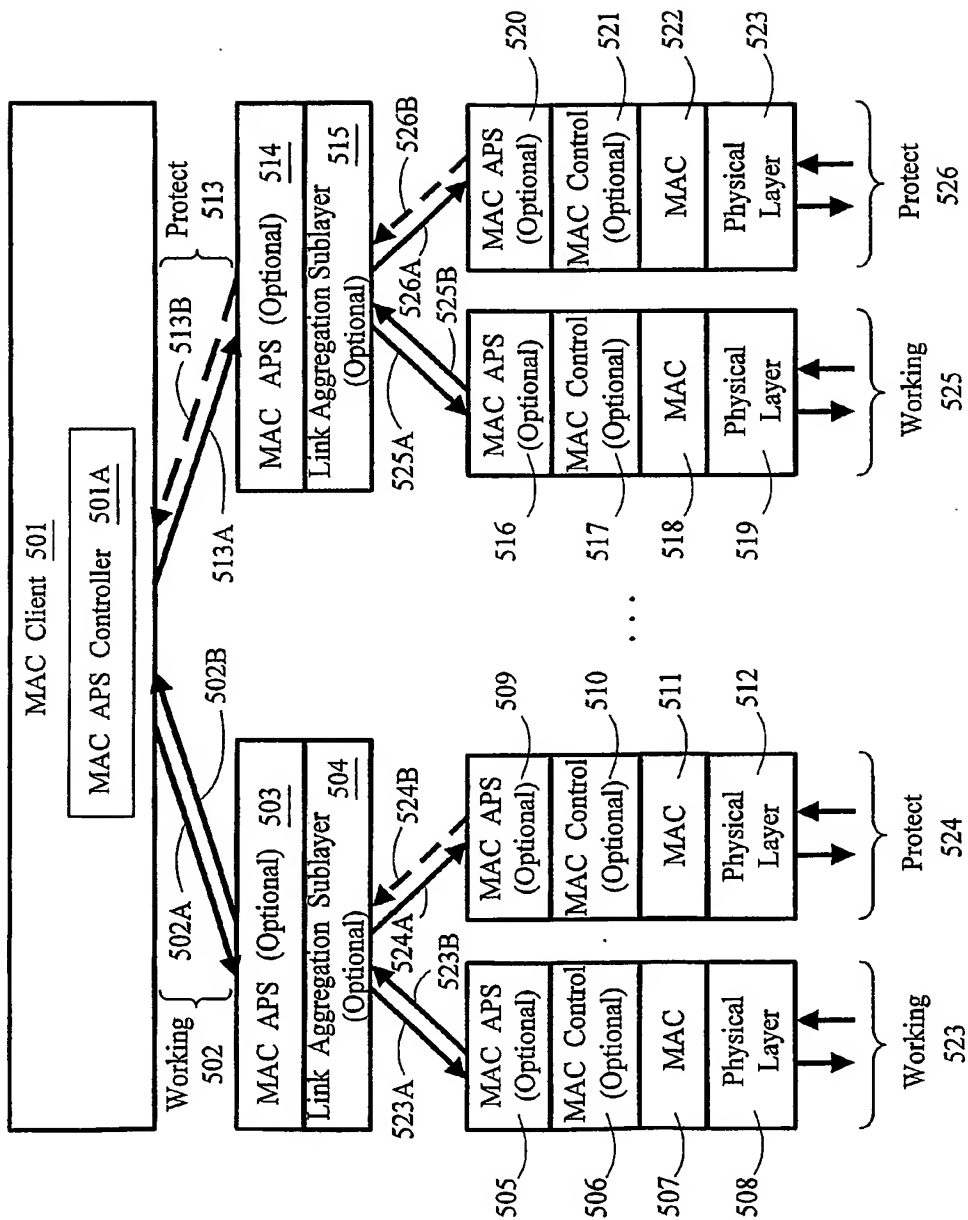


FIG. 5

6/6

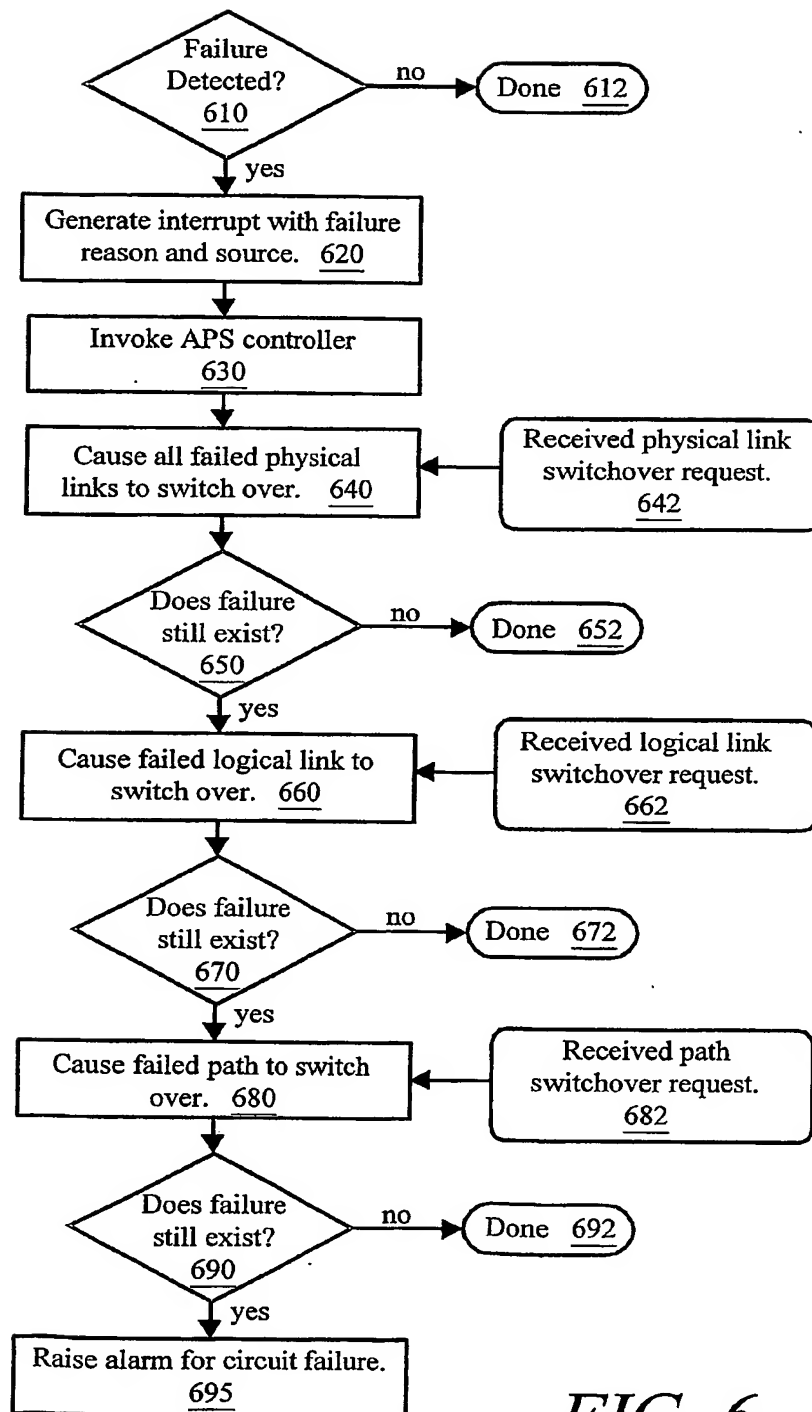


FIG. 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/28112

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : H04L 1/22
US CL : 370/228

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 370/228

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	US 6,359,858 B1 (SMITH et al) 19 March 2002. See columns 8 and 9.	1,16,17,20,21,23,28,3 2,34
A	US 6,188,667 B1 (ROBERTS et al) 13 February 2001. See entire document.	1-39
A	US 5,623,357 A (KIGHT et al) 22 April 1997. See entire document.	1-39
A	US 5,740,157 A (DEMIRAY et al) 14 April 1998. See entire document.	1-39
A	US 5,671,213 A (KURANO) 23 September 1997. See entire document.	1-39

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

Inter document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

23 October 2002 (23.10.2002)

Date of mailing of the international search report

06 DEC 2002

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Authorized officer

Hassan Kizou

Facsimile No.

Telephone No. 703-305-4150